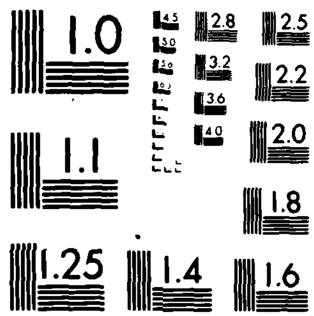
AD-A087 649   CATHOLIC UNIV OF AMERICA WASHINGTON D C HUMAN PERFOR--ETC F/G 5/8
SEQUENTIAL STRUCTURE AND CONTEXT IN THE CLASSIFICATION OF NONSP--ETC(U)
JUL 80 J H HOWARD, J A BALLAS   N00014-79-C-0550

UNCLASSIFIED   TR-80-14-ONR   NL

1 of 1

END
DATE FILMED
9 80
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

Sequential Structure and Context in the Classification of Nonspeech Transient Patterns

James H. Howard, Jr. and James A. Ballas

ONR CONTRACT NUMBER N00014-79-C-0550

Technical Report ONR-80-14

Human Performance Laboratory
The Catholic University of America

July, 1980

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ONR-80-14

2. GOVT ACCESSION NO.
AD-A087

3. RECIPIENT'S CATALOG NUMBER
649

4. TITLE (and Subtitle)
(6) Sequential Structure and Context in the Classification of Nonspeech Transient Patterns,

5. TYPE OF REPORT & PERIOD COVERED
(9) Technical Rep[redacted]t,

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
(10) James H. Howard, Jr. [redacted] James A. Ballas

8. CONTRACT OR GRANT NUMBER(s)
(15) N00014-79-C-0550

9. PERFORMING ORGANIZATION NAME AND ADDRESS
The Catholic University of America
Washington, D.C. 20064

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
NR 196-159 (12) +3

11. CONTROLLING OFFICE NAME AND ADDRESS
Engineering Psychology Programs, Code 455
Office of Naval Research
Arlington, Virginia 22217

12. REPORT DATE
(11) 15 Jul[redacted]80

13. NUMBER OF PAGES
35

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
(14) TR-12-24-ONR

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
auditory classification
nonspeech transient recognition
top-down processes
semantic context
sequential pattern structure

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Three experiments were conducted to investigate the role of both syntactic (i.e., sequential structure) and semantic (i.e., contextual knowledge of the source events) factors in a two-alternative (target/nontarget) categorization task involving patterns of nonspeech acoustic transients. The results demonstrated that both factors can play an important role in the classification of such patterns. Although pattern syntax influenced performance in all three experiments, the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

409387

effects of syntactic structure were clearest in Experiment 1 in which listeners categorized meaningless tonal patterns. Listeners who categorized a syntactically structured target set performed better than those with an unstructured set. Experiments 2 and 3 were similar to Experiment 1, but listeners classified patterns of familiar, brief-duration, complex sounds rather than tones. When listeners in Experiment 3 were given explicit descriptive information about the pattern components in their instructions, performance actually improved for interpretable patterns but was slightly degraded for uninterpretable patterns. This suggests that syntactic and semantic factors interact in an important way to influence performance. It was argued that many complex nonspeech patterns have both syntactic and semantic structure which is determined by the sequence of source events which produce them. In classifying such patterns, as in the case of speech, listeners rely on their knowledge of these factors as well as on the perceptual information in the sound itself.

Sequential Structure and Context in the Classification of Nonspeech Transient Patterns

Many current theorists have argued that the recognition of fluent speech involves both top-down or knowledge driven and bottom-up or data driven processes (Marslen-Wilson & Welsh, 1978; Marslen-Wilson & Tyler, 1980; Cole & Jakimik, 1978, 1980). In other words, when human listeners perceive speech they appear to use their general knowledge of linguistic structure (both syntactic and semantic) as well as the specific perceptual information in the signal. In contrast, relatively little research has investigated the role of syntactic and semantic factors in the perception of complex nonspeech patterns. Although less obvious than the case of speech, many of the complex nonspeech sounds which we recognize in everyday life have a specifiable sequential structure (syntax) as well as a semantic context. The importance of sound effects in radio drama confirms this point. One sequence of acoustic transients is heard as someone opening a door to enter a room, whereas another can depict the escape of bank robbers with the police in hot pursuit. Ordered sequences of nonspeech transients of this sort will be referred to as transient patterns. These patterns have sequential structure since their individual components occur in an order and with durations determined by the source events. For many patterns, an experienced listener can identify the source events when presented with only the sound pattern. The present paper investigates the role of syntactic (i.e.,

sequential structure) and semantic (i.e., contextual knowledge of the source events) factors in transient pattern recognition.

The most convincing evidence that aural perception involves both bottom-up and top-down processing is found in the speech perception literature. Logically, the "raw data" or specific sounds in continuous speech cannot be sufficient to account for language understanding since the "raw" input is neither complete nor unambiguous (Cole & Jakimik, 1978). Rather, the listener must rely on his or her knowledge of the syntax and semantics of language, and the constraints introduced by those elements that can be interpreted unambiguously. It is not uncommon for us to "hear" missing words or to correct mispronounced words when they occur in fluent speech. For example, Warren (1970) has demonstrated a "phoneme restoration" effect in the perception of spoken text. His listeners consistently reported hearing phonemes that had actually been replaced by a buzz or other nonspeech sound. Warren argued that this reflects the operation of a higher-level process that perceptually produces the missing phoneme.

Similarly, Marslen-Wilson and Welsh (1978) investigated the tendency for listeners to correct mispronounced words while shadowing continuous text. They concluded that although speech perception in this context is primarily data-driven, top-down processes serve to make the system more resistant to input noise and to enhance overall recognition efficiency. In other papers, Cole and Jakimik (1978, 1980) described experiments which investigated a variety of factors in word recognition. They

concluded that listeners use a number of knowledge sources in speech perception ranging from fairly specific item-to-item syntactic constraints to global semantic considerations such as the theme or title of a story. In conclusion they argued "...it is not only what we hear that tells us what we know; what we know tells us what we hear" (1978, p. 113).

The evidence from the speech literature that top-down processes play a major role in perception is not particularly surprising. Less obvious, however, is the evidence reported in recent years by Bregman and his associates (Bregman, 1978) which demonstrates a parallel role of knowledge-driven processes in the perception of relatively simple and semantically impoverished tonal sequences.

Bregman's basic assumption is that listeners are "built to pay attention to acoustic sources, not to acoustic components..." (p. 74). At any point in time the single waveform we hear is likely to represent information combined from several sources, and yet we perceive sounds from each of the separate sources individually (the dog barking, a Bach cantata, the telephone ringing, for example) rather than a nonsensical hodgepodge. Bregman has referred to this perceptual phenomenon--the act of sorting our perceived acoustic world into separate sources--as auditory streaming. From a theoretical vantage he has argued that streaming occurs as the result of a perceptual parsing of the complex acoustic input (Bregman, 1978), and his research has focused on identifying the rules involved in the formation of auditory streams.

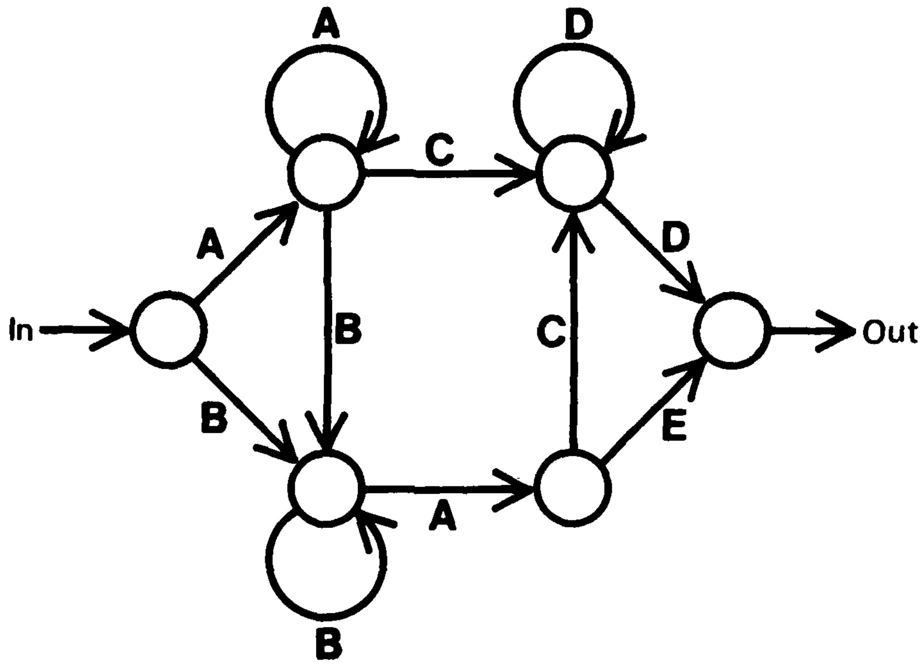
To date auditory streaming research has been restricted largely to relatively simple stimulus contexts. For example, a listener may be asked to judge whether one or two melodic passages are heard when six low-frequency pure tones are played alternately with six high-frequency pure tones. Two separate streams or patterns are readily heard in this context, and the influence of factors such as tonal frequency separation, duration, and alternation rate on stream formation have been investigated. The results of these experiments have been consistent with similar research on simple visual patterns in revealing general heuristics related to the classical Gestalt principles. Such things as similarity, good continuation, simplicity, common fate and closure all operate to influence the auditory streams that will be formed in a specific context.

In general, Bregman (1978) has concluded that two kinds of factors are important in auditory streaming. First, there are general--and probably innate--rules or heuristics that can be applied to parse all signals. Second, there are more specific rules related to a listener's skills, intentions, and knowledge of the stimuli that apply to only selected patterns. By restricting his research to relatively simple acoustic patterns, Bregman has necessarily concentrated on the former of these factors. In the present study we investigate more complex transient patterns and consequently extend Bregman's analysis to heuristics of the second type.

Our approach to this problem involves an extension of Reber's "implicit learning" procedure. In his research, Reber

(Reber, 1969, 1976; Reber & Allen, 1978; Reber & Lewis, 1977) examined the role of syntactic structure in the classification of visually presented letter patterns. Listeners classified either grammatical or nongrammatical patterns. The grammatical patterns were generated by a simple finite-state grammar similar to the one shown in the state-transition diagram of Figure 1. An additional letter in the pattern string is produced with every legal state-transition made between the initial and terminal states. For example, the letter pattern "AAACDD" could be produced by the grammar and is therefore grammatical, whereas the pattern "AADDCC" would be ungrammatical since it could not be produced by the grammar.

Reber's extensive research has demonstrated consistently superior classification performance with the grammatical as opposed to nongrammatical, arbitrarily grouped patterns. This has been shown using a variety of tasks and dependent variables. In the present study, three experiments were conducted to investigate the role of syntactic structure in a two-alternative transient pattern classification task. The simple finite-state grammar of Figure 1 was used to generate syntactically structured patterns for all three experiments. In the first experiment listeners classified meaningless patterns of brief-duration pure tones. The second experiment was similar to the first, but the pattern components consisted of complex familiar sounds rather than simple tones. Although the individual components were familiar in this experiment, the transient <u>patterns</u> were not interpretable. Finally, the role of



Grammatical Codes	Experimental Stimuli		
	Exp 1	Exp 2	Exp 3
A	1157 Hz	Drill	Valve
B	1250 Hz	Clap	Drop
C	1345 Hz	Steam	Steam
D	1442 Hz	Clank	Clang
E	1542 Hz	Wood	Flush

Figure 1. State-transition diagram for the finite-state grammar used to generate structured target patterns in all three experiments.

both syntactic and semantic factors was investigated in the third experiment in which some listeners classified semantically interpretable patterns of complex sounds.

Experiment 1

The first experiment was designed to demonstrate that listeners can use syntactic information to facilitate the classification of nonspeech transient patterns. In Experiment 1 listeners were required to classify sequences of brief-duration pure tones as either "target" or "noise" patterns. For the Grammatical group the target patterns were generated by the finite-state grammar of Figure 1, whereas for the other, Nongrammatical group the targets were randomly determined but matched to the grammatical targets in length. By comparing performance with structured (Grammatical group) and unstructured (Nongrammatical group) target patterns, we can assess the importance of syntactic or sequential structure in the classification of simple unfamiliar tonal patterns.

Method and Procedure

Participants. Ten student volunteers served as listeners in the experiment. Five were assigned to each of the two groups.

Stimuli. Individual transient events consisted of five pure tones selected to be approximately equally spaced in pitch (1157, 1250, 1345, 1442, and 1542 Hz). The grammatical patterns were produced by assigning one of the five tones to each of the

output letters shown in Figure 1 in corresponding ascending order. Twelve grammatical patterns ranging in length from four to six events (three, four, and five patterns of each length, respectively) were selected to make up the "grammatical target" category. A corresponding "nongrammatical target" set was produced by randomly sampling patterns from the total set of possible patterns with the restriction that sampled targets match those of the grammatical target patterns in length. Similarly, 48 randomly constructed "noise" patterns were selected to be non-overlapping with the target sets, but to match them in length. Within the patterns each tone was presented for 80 msec at a comfortable listening level (87 dB SPL). Successive tones were separated by 20 msec of silence.

Apparatus. All experimental events were controlled by a general-purpose laboratory computer. The tones were synthesized with the computer using standard digital techniques. They were output on a 12-bit digital-to-analog converter at a sampling rate of 12.5 kHz, low-pass filtered at 5 kHz (Khron-Hite model 3550), attenuated, and presented binaurally over matched Telephonics TDH-49 headphones with MX-41/AR cushions. Verbal prompts were presented on a video monitor in the testing booth, and listeners indicated their responses by pressing buttons on a solid state keyboard.

Procedure. Listeners were tested individually in a sound attenuated booth. The experiment began when the listeners were instructed that they would be hearing patterns made up of several notes played very quickly. They were told that some of

the patterns were designated as targets and that their task would be to pick out the targets. Although listeners were told that targets and nontargets could occur equally often, no information was provided regarding the composition of the target set. However, they were told that the pattern categories were determined by the order of components and that loudness and duration were not relevant to the classification. The Grammatical and Nongrammatical groups received identical instructions.

Each trial began when the word "LISTEN" appeared on the listener's screen. A second prompt, "TARGET (Y OR N)?" followed the pattern presentation. The listener then responded by pressing "Y" or "N" on the keypad, and visual feedback was provided immediately after the response. After a brief inter-trial-interval (1.5 sec), the screen was erased and the next trial began. Each listener received 96 trials (four presentations of each of the 12 targets and 48 presentations of nontargets) in each of 12 blocks. Pattern presentation order was randomized within blocks, and listeners completed four blocks on each of three days. Overall, there were 1152 trials per individual.

Immediately after the last block, listeners in both groups were told that we had used a set of rules--like the rules of language--to construct the target patterns. We explained that they would be hearing a new set of patterns and that their task would be to classify each pattern as target or nontarget, "just as you can tell if a sentence is grammatically correct without

knowing all the rules for sentences, so should you be able to tell whether any sound is consistent with the rules we used by remembering how the targets sounded." They then completed an additional block of 96 trials responding as before, but without feedback. The target sounds in this test block were selected from the grammatical patterns produced by the grammar in Figure 1 which were not used as targets in the experiment. This test condition was included to determine whether the listeners in the Grammatical group could use their syntactic knowledge to classify novel, but grammatical, patterns. Each listener was interviewed before leaving.

Results and Discussion

The hit (responding "yes" to a target) and false alarm (responding "yes" to a nontarget) rates were used to compute a response bias free (d') index of performance for each individual on each block. These data were then averaged across listeners within each of the two groups to assess group performance. Nearly perfect performance (hit rate of .99, false alarm rate of .01) would yield a d' of 4.64. These results are displayed in Figure 2.

Although the data are not strictly monotonic over blocks, it is clear that performance improved with practice for both groups. This effect was revealed statistically by a significant Block effect in a two-way (Group by Block) mixed-design analysis of variance with repeated measures on the Block factor ($\underline{F}(11,88) = 14.91$, $\underline{p} < .001$).

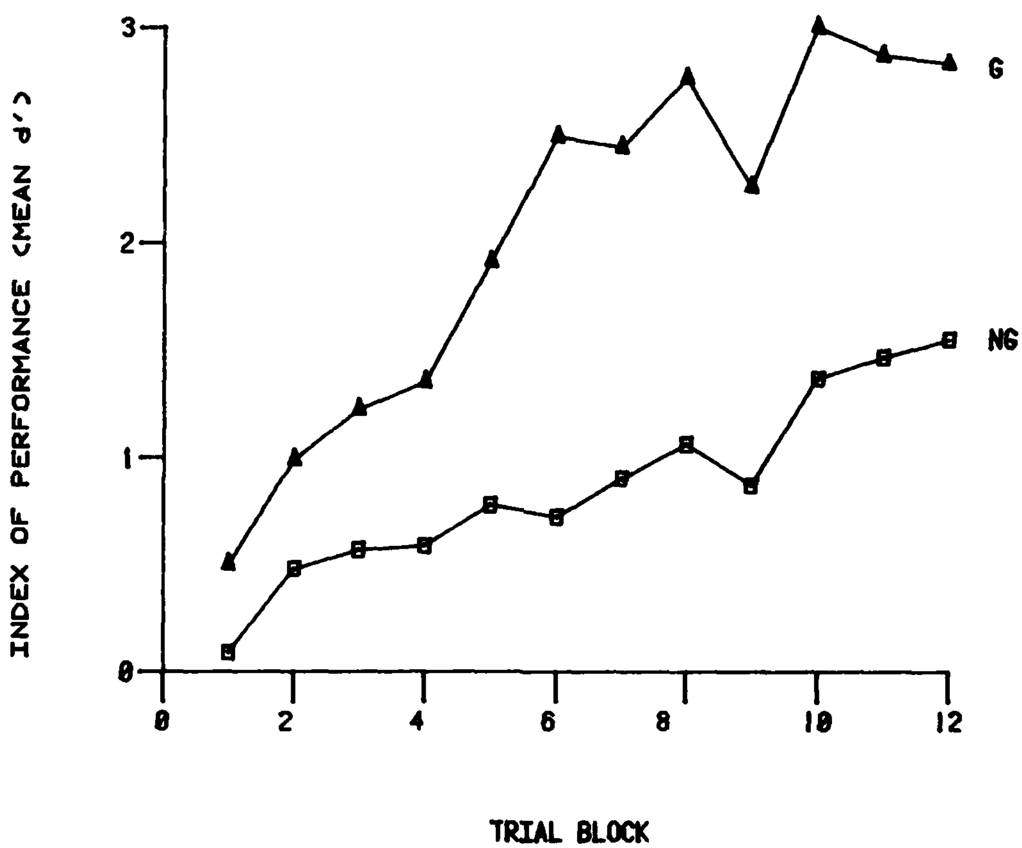


Figure 2. Mean performance on classification of brief-duration pure tones for structured (G) and unstructured (NG) patterns.

It is also clear from Figure 2 that the Grammatical group reached a substantially higher overall performance level than did the Nongrammatical group. This finding was demonstrated by a significant Block by Group interaction in the analysis of variance ($\underline{F}(11,88) = 2.23$, $\underline{p} < .025$). The main effect of group was only marginally significant ($\underline{F}(1,8) = 4.27$, $\underline{p} < .10$) indicating that the effect of grammaticality developed primarily with practice. Overall, these findings are consistent with Reber's earlier results with letter strings (Reber, 1969) and with our hypothesis that listeners can use syntactic structure to help them classify complex nonspeech patterns.

Of further interest is the performance of listeners in the Grammatical group with unfamiliar patterns in the final post-test block. If listeners in the Grammatical group had internalized the syntactic or grammatical structure of the target patterns during the experiment, then their performance should be substantially better than chance on the test block. On the other hand, listeners in the Nongrammatical group would have had no opportunity to learn about the pattern grammar and consequently their performance should be considerably worse than that of the Grammatical group. A d' index of performance was computed for each of the ten listeners on the final test block. These data are presented in Table 1. It is evident from these data that listeners in the Grammatical group performed substantially better than the near chance levels (i.e., d' = 0.0) of listeners in the Nongrammatical group. This difference in group mean performance was found to be

Table 1

Final Test Block Performance (d') for All Listeners
in Each of the Three Experiments

	Group	Individual Listener					
		1	2	3	4	5	Mean
Experiment 1	G	-1.48	1.48	1.84	2.58	3.31	1.55
	NG	-.66	-.43	-.23	.10	.56	-.13
Experiment 2	G	-.34	.10	.32	1.76	3.79	1.13
	NG	-.40	-.12	-.10	.10	.20	-.06
Experiment 3	G/S	.95	1.26	1.48	2.35	____	1.51
	G/NS	.86	1.23	1.95	2.55	____	1.65
	NG/S	-1.20	.00	.00	.34	____	-.21
	NG/NS	-.88	-.10	.30	1.12	____	.11

statistically reliable ($t(8) = 1.98$, $p < .05$, one-tailed). This difference occurred despite a relatively large negative d' value observed for listener 1 in the Grammatical group. The large absolute magnitude of this value indicates that this individual was able to distinguish the grammatical and nongrammatical patterns to some extent, but had classified the targets as nontargets and vice versa. Since no feedback was provided during the test block, a response reversal of this sort would not be detected easily by the listener. Overall, performance on the final block supports our position that listeners in the Grammatical group had actually learned something about the syntactic rules used to generate the target patterns they had classified previously.

Experiment 2

Experiment 1 demonstrated that listeners can use syntactic pattern structure to their advantage in classifying simple tonal transient patterns. The question arises as to whether a similar result would occur for sound patterns made up of complex, realistic transients. To investigate this, a second experiment was conducted in which listeners classified patterns as either targets or nontargets under conditions similar to those of Experiment 1. One group of listeners had targets produced by the grammar of Figure 1, whereas the other group had a randomly constructed target set. However, unlike the previous experiment, the individual transient events used here were familiar, but unrelated real world sounds recorded in the

laboratory.

Method

Participants. Ten student volunteers served in the experiment, five in the Grammatical group, and five in the Nongrammatical group. None had served in the previous experiment.

Stimuli. Five individual acoustic transients were selected from a larger set of common "real world" sounds collected in our laboratory. The larger set was produced by recording a variety of events such as a "clank" (hammer striking a heavy metal object), a "thump" (a hollow, resonant sound from striking a metal drum), and other similar sounds. These samples were then digitized using standard signal processing techniques with a 10-bit analog-to-digital converter at a 12.5 kHz sampling rate. The initial 82 msec segment of each sound was then submitted to a physical analysis before being used in the experiment. A name and brief description for each transient is presented in Table 2.

The 12 grammatical targets, 12 nongrammatical targets, and 48 noise patterns were generated as in Experiment 1. Each sound was presented for 82 msec at a comfortable listening level that differed for each sound. Successive transients were separated by 510 msec of silence within a pattern.

Apparatus. Same as in Experiment 1.

Procedure. The procedure was identical to that of Experiment 1.

Table 2

Transient Sounds used in Experiments 2 and 3

Experiment 2	
Drill	82 msec recording of a high speed hand drill being turned on
Clap	82 msec recording of a hand clap
Steam	82 msec recording of white noise bandpass filtered between 4.6 kHz and 5.4 kHz
Clank	82 msec recording of a hammer striking a C-clamp
Wood	82 msec recording of two pieces of wood being struck together
Experiment 3	
Open Valve	320 msec recording of a radiator valve being turned
Water Drop	38 msec recording of a drop of water
Steam	320 msec recording of white noise bandpass filtered between 4.6 kHz and 5.4 kHz
Clang	320 msec recording of a metal object striking a radiator pipe
Water Flush	320 msec recording of water flushing down a drain

Results and Discussion

Hit and false alarm rates were used to compute a response bias free (d') index of performance for each individual on each block. These data were then averaged across listeners within the two groups to assess group performance. These data are presented in Figure 3. Overall, the results are similar to those of Experiment 1. Although it appears that performance for the Nongrammatical group may have reached an asymptote by block 9, performance generally improved with practice for both groups as indicated by a significant Block effect in a two-way (Group by Block) mixed-design analysis of variance ($F(11,88) = 6.95$, $p < .001$). Furthermore, the Grammatical group did better than the Nongrammatical group on all but the first block. As in Experiment 1, however, the performance difference between the two groups developed with practice since the Group factor interacted reliably with Block ($F(11,88) = 2.82$, $p < .005$), but did not produce a significant main effect ($F(1,8) = 2.79$, $p > .10$). On the basis of these findings we can generalize the conclusion of Experiment 1 to include patterns of familiar transients as well as patterns of simple tones. It appears that listeners are able to use syntactic or grammatical structure to facilitate classification in both cases.

However, it is interesting to note that the overall performance level reached by the Grammatical group in the present experiment (mean d' = 1.51) was considerably lower than that reached in the earlier, tonal pattern experiment (mean d' = 2.06). Furthermore, this difference cannot be attributed

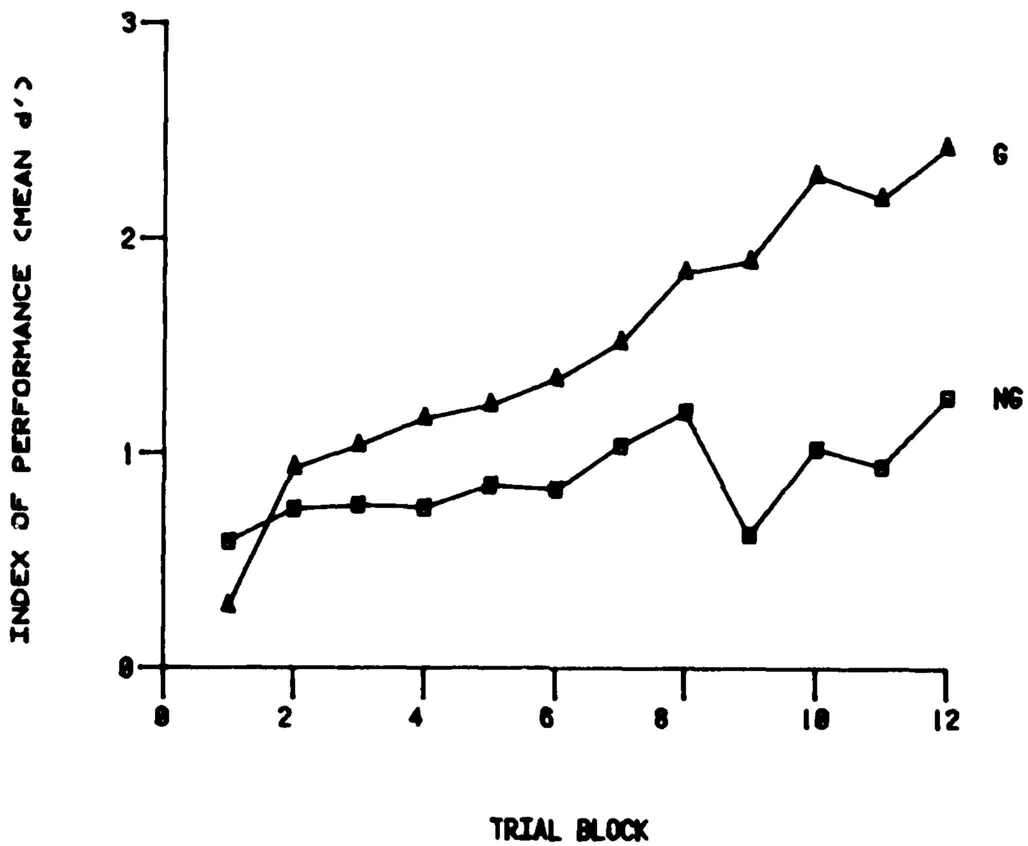


Figure 3. Mean performance on classification of "real world" sounds for structured (G) and unstructured (NG) patterns.

to the different pattern components used since the mean performance levels observed for the Nongrammatical groups were virtually identical across the two experiments (mean d' = .87, Experiment 1; mean d' = .88, Experiment 2). This suggests that structured patterns of unrelated, familiar sounds may be more difficult for listeners to classify than structured patterns of simple tones.

Although the present experiment was designed to investigate the role of syntactic processes in pattern classification, the overall difference observed between Experiments 1 and 2 can be explained most easily by referring to semantic processing. In the present experiment, the listeners were able to recognize the individual transients. In the words of one participant, "These are the kinds of sounds we walk around all day trying to ignore." Since the sounds were familiar, the listener could not avoid using a parsing strategy that tried to make "sense" out of the patterns. Since the finite-state grammar we used was semantically arbitrary, this would prove to be an impossible task. On the other hand, Grammatical listeners in Experiment 1 had little difficulty in using purely syntactic parsing rules since they did not expect the tonal transients to form semantically sensible patterns. This point is dramatized by one outspoken listener in Experiment 2 who offered the unsolicited advice that we should have used tones instead of sounds to make the task easier!

Finally, the results obtained for the final test block with novel grammatical patterns revealed generally poor performance

for all listeners with only two individuals in the Grammatical group (listeners 4 and 5) performing better than chance. The performance level observed for each individual is shown in Table 1. Although the group mean performance was somewhat greater for the Grammatical group than the Nongrammatical group, this difference did not approach significance ($\underline{t}(8) = .99$). This finding is consistent with the above discussion in suggesting that the listeners have a great deal of difficulty in abstracting syntactic structure from patterns of familiar sounds that do not relate to any interpretable sequence of source events. The two listeners in the Grammatical group who seemed able to do this may have ignored the distracting semantic contents of the individual pattern components successfully.

Experiment 3

The results of the two experiments reported above suggest that a more elaborate investigation of both syntactic and semantic factors in transient classification would be appropriate. In Experiments 1 and 2 we found that syntactically structured patterns were easier to classify than unstructured patterns. Listeners in the present experiment were also required to classify either grammatical or nongrammatical target patterns, but the procedure was extended so that we could evaluate explicitly the semantic component in aural classification.

A comparison of our findings from Experiments 1 and 2 suggested that the semantic cues provided by individual

transient components may be distracting when the overall pattern does not lend itself to an interpretable semantic analysis. In the present experiment, some listeners were required to classify transient patterns that were semantically sensible. To include this condition it is essential to have patterns that are both syntactically and semantically reasonable. In other words, the grammar cannot be arbitrary, but must reflect the temporal structure of possible real-world events. Similarly, the selection of individual transient events must be consistent with the grammar.

The simple finite-state grammar of Figure 1 had been developed with these criteria in mind. The grammar can represent possible temporal relations among a series of water and steam related events when appropriate complex transients are substituted for the tones and sounds used in the preceding experiments. The five sounds employed in Experiment 3 are described briefly in Table 2. To illustrate how semantically interpretable patterns can be produced, consider an output string A-A-A-C-D-D from the grammar in Figure 1. This corresponds to a pattern that could represent someone taking three turns to open a valve which releases steam which, in turn, causes pipes to clang twice. Similar source scenarios can be provided for other grammatical patterns. To evaluate the possible role of semantic context in aural transient classification, one half of the listeners in the present study were read a brief paragraph that suggested a schema or theme for the patterns they would hear. The paragraph was suggestive, but

did not identify any specific patterns explicitly:

> All of the individual sounds relate to water and steam. You will hear such things as drips, water flushing down a drain, a valve being turned on, steam escaping, and radiator pipes clanging.

The remaining listeners received no semantic information about the patterns. The role of semantic context in transient classification can be assessed by comparing performance across the two instructional conditions.

To summarize, four groups were tested in the present experiment. The groups were determined by factorially combining the two syntactic (Grammatical and Nongrammatical) and two semantic (Semantic instructions and No Semantic instructions) variables. The Grammatical/Semantic group classified structured target patterns and received the semantic information described above, whereas the Grammatical/No Semantic group categorized the same structured target patterns without any explicit semantic instructions. Two corresponding nongrammatical target groups were tested (Nongrammatical/Semantic and Nongrammatical/No Semantic). The possible interaction of the syntactic and semantic factors was of particular interest here. Specifically, the findings discussed above suggest that explicit semantic instructions may induce a semantic parsing strategy which would facilitate classification for interpretable patterns (Grammatical group), but interfere with performance when the patterns were not interpretable (Nongrammatical group).

Method

Participants. Sixteen student volunteers served as listeners in the experiment, four in each group. No listener had served in either of the previous experiments.

Stimuli. The five thematically related transient sounds described in Table 2 were recorded and digitized as in Experiment 2. These were then combined to form sequential transient patterns as in the earlier experiments. Twelve grammatical target patterns were produced using the grammar of Figure 1. The unstructured target and noise patterns were constructed randomly as in the earlier experiments. Within a pattern, each transient was presented for a brief duration (32 msec for the drip and approximately 300 msec for all others) at a comfortable listening level that differed slightly for the various sounds to enhance realism. Successive sounds were separated by 510 msec within the patterns.

Apparatus. Same as in Experiments 1 and 2.

Procedure. The procedure was identical to that of Experiments 1 and 2. All four groups were tested for 12 blocks on the target patterns with feedback and for one block on the test patterns without feedback.

Results and Discussion

The hit and false alarm rates were used to compute a response bias free (d') index of performance for each individual on each block. These data were then collapsed across individuals within each group to determine group performance

levels. These mean data are plotted across blocks for each of the four groups in Figure 4. This finding was confirmed statistically by a significant main effect of Block in a three-way (Pattern Structure by Semantic Instructions by Block), mixed-design analysis of variance ($F(11,132) = 25.33$, $p < .001$). It is clear from Figure 4, however, that large differences exist in the effects of practice across the four groups. In particular, the two Grammatical groups showed considerable improvement with practice, whereas the two Nongrammatical groups showed relatively little improvement. This was revealed by a statistically significant Pattern Structure by Block interaction ($F(11,132) = 9.96$, $p < .001$) and is consistent with the results of Experiments 1 and 2. No interaction was observed between the Semantic Instruction and Block factors ($F(11,132) < 1.0$).

Second, listeners in the two Grammatical groups also performed at a significantly higher level than did listeners in the two Nongrammatical groups. This was supported by a significant main effect of Pattern Structure ($F(1,12) = 59.11$, $p < .001$). This finding is not consistent with the earlier experiments in which the effect of pattern syntax emerged largely with practice. It suggests that the facilitative effect of sequential structure was stronger in the present experiment than in the earlier studies.

Third, there appears to be no overall difference between listeners who were provided with a semantic context and those who were not. The main effect of Semantic Instruction did not approach statistical significance ($F(1,12) < 1.0$).

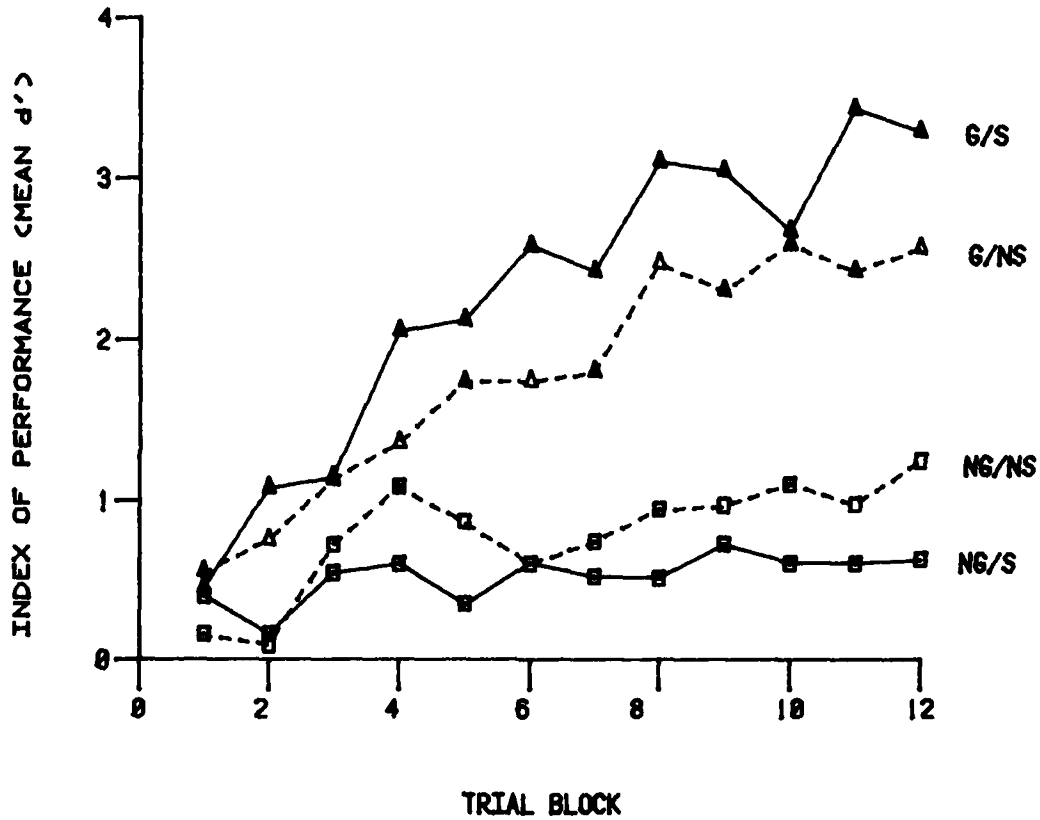


Figure 4. Mean performance on classification of "real world" sounds for structured (G) and unstructured (NG) patterns, with (S) and without (NS) thematic information about the sounds.

Nevertheless, it is obvious that the semantic instructions influenced classification performance. Specifically, semantic instructions appeared to enhance performance for those listeners who also received syntactically structured target patterns (Grammatical/Semantic vs. Grammatical/No Semantic), but impair performance for those who classified the syntactically unstructured targets (Nongrammatical/No Semantic vs. Nongrammatical/Semantic). This result was seen in a statistically reliable Pattern Syntax by Semantic Instruction interaction ($F(1,12) = 4.73$; 4.75 required for $p = .05$). This effect was examined further in a post-hoc analysis. Lindquist's test of critical differences revealed that the Grammatical/Semantic group performed significantly better than the Grammatical/No Semantic group (observed difference of .51; critical difference of .51 for $p = .05$). However, the apparent difference between the Nongrammatical/Semantic and Nongrammatical/No Semantic groups did not reach statistical significance (observed difference of .27; critical difference of .51 for $p = .05$).

The above finding clearly demonstrates the importance of semantic context in aural transient classification. In addition, it underscores our earlier conclusion that these effects are not always facilitative. In particular, semantic cues--in our case the explicit semantic description of the sounds--appear to induce semantic parsing strategies that can enhance performance only for semantically interpretable patterns (Grammatical groups). Listeners in the Grammatical/Semantic

group were able to use the semantic information we provided to their advantage. Here, a semantic parsing strategy was appropriate in that it could lead to sensible interpretations for the target patterns. These listeners performed reliably better than their counterparts who had to depend on the syntactic structure of the patterns and the implicit semantic cues in the isolated transients alone.

On the other hand, the semantic instructions led to a slight, albeit nonsignificant, impairment of performance when semantically anomalous patterns were used (Nongrammatical groups). For these individuals, the specific thematic instructions inappropriately led them to search for sensible interpretations of the patterns when none existed.

Although the explicit thematic instructions were an obvious source of semantic information in the present study, it is also obvious that the familiar sounds themselves provided an additional source of semantic cues. Since these cues were available for listeners in all four groups, it was not possible to assess their effects explicitly in the present study.

Finally, the final test block performance was examined to determine whether any listeners were able to generalize their knowledge of the target set to new, grammatical test patterns. The d' performance levels on the final test block are shown for each individual in Table 1. As expected, listeners in the two Nongrammatical groups responded at approximately chance levels, whereas listeners in both Grammatical groups responded at above chance levels. Overall, listeners in the two Grammatical groups

(mean d' = 1.58) performed reliably better on the test patterns than did listeners in the two Nongrammatical groups (mean d' = -.05) ($t(14) = 2.40$, $p < .025$, one-tailed). Those who received semantic instructions performed slightly worse than those with no explicit semantic instructions, but this difference was not statistically reliable ($t(6) = .21$). These findings indicate that listeners in the Grammatical groups were able to internalize aspects of the pattern grammar regardless of whether explicit semantic instructions were provided.

General Discussion

Overall, the results presented above have demonstrated that both syntactic and semantic factors can play an important role in the classification of acoustic transient patterns. Although pattern syntax influenced performance in all three experiments, the effects of syntactic structure were most clearly seen in Experiment 1 in which listeners categorized meaningless tonal patterns. Here, listeners who categorized a grammatically structured target set performed substantially better than those with an unstructured set. Listeners in the former group were also able to generalize their knowledge of the grammar to a novel set of grammatical test patterns. These results are consistent with Reber's earlier findings with visually-presented, meaningless letter strings.

Reber has argued that listeners exposed to structured stimuli internalize a "conceptual structure" which represents the underlying grammar or rules, and that this abstraction

process occurs implicitly rather than explicitly. In this sense he has argued that the learning of synthetic grammars of the sort employed in the present study is similar to the acquisition of natural language grammars (Reber & Allen, 1978). Our findings are generally consistent with this interpretation since in the post-experimental interview, listeners found it impossible to articulate the rules they used to identify the target patterns. Although a few listeners were able to indicate some obvious properties of the grammatical patterns (e.g., the fact that they began with one of two sounds), most specified such vague classification rules as: the targets were "more coherent" or "flowed better" and the nontargets were "unexplainably different" or "not harmonious." Regardless of how the pattern structure is internalized, however, the present results make it clear that syntactic structure does influence the processing of complex transient patterns.

Furthermore, it is obvious from Experiments 2 and 3 that the effects of pattern syntax cannot be considered in isolation. Rather, syntactic and semantic factors interact in an important way to determine categorization performance. For example, in Experiment 2 listeners categorized uninterpretable patterns of familiar sounds. Although clear syntactic effects were observed in this experiment, the effect of pattern structure was considerably smaller than the corresponding effect in Experiment 1. This suggests that the listeners' semantic knowledge (i.e., their familiarity with the pattern components) actually interfered with their ability to abstract the pattern structure.

The importance of semantic factors in aural classification was even more obvious in the third experiment. When listeners were given explicit descriptive information about the pattern components in their instructions, performance actually improved for interpretable patterns but was slightly degraded for uninterpretable patterns. Although Cole and Jakimik (1978) have demonstrated that the theme or title of a story influences the linguistic processing of specific words, the strength of the present effect with nonspeech patterns is somewhat surprising.

One explanation of the effect is based on a relatively simple labeling strategy. It is possible that the descriptive instructions we provided equipped the listeners with a consistent set of labels for the pattern components. As a result, these listeners could employ more effective encoding and chunking strategies to facilitate the learning of pattern/category pairs in a paired associate fashion. While it appears likely that labeling differences of this sort played a role in Experiment 3, it is apparent that the semantic instruction effect cannot be attributed exclusively to labeling. In particular, this explanation cannot account for the absence of labeling facilitation for the uninterpretable patterns (the Nongrammatical/Semantic condition).

A more compelling explanation proposes that listeners are influenced by existing semantic structures when perceiving patterns of familiar complex sounds. These existing structures have been referred to as <u>frames</u> (Minsky, 1975) or <u>scripts</u> (Schank & Abelson, 1977). In Minsky's view, a frame is simply a

"data-structure for representing a stereotyped situation" (1975, p. 212). We propose that most individuals have frames for a wide range of possible acoustic transient patterns. Each frame represents the source events for a particular pattern. When an initial sound occurs, the listener refers to the likely source scenarios--the frames--which contain the sound as a beginning component. In other words, the listener constructs hypotheses about what the entire pattern will be, based on partial perceptual information and his or her existing knowledge. As successive transients are heard and interpreted, inappropriate frames can be eliminated until, ultimately, enough information is accumulated for the pattern to be associated with an appropriate source scenario. In this view, the interpretation of complex transient patterns results from an interplay of bottom-up and top-down processes.

In such a system, explicit semantic instructions would not only provide the listener with a set of component labels, but with a set of possible frames as well. In Experiment 3, the instructed listeners would attempt to relate the patterns they heard to familiar scenarios involving steam and water flow. These frames or scenarios would be appropriate in the case of interpretable (i.e., grammatical) patterns, but inappropriate for the uninterpretable patterns. In the latter case, the listeners' inability to interpret the patterns using the suggested frames would prove distracting. In other words, these listeners would be unable to make semantic sense out of the patterns despite the fact that the individual pattern components

were familiar and consistent with the labels we provided. The fact that an explicit semantic context led to degraded peformance with uninterpretable patterns suggests that it may be very difficult to ignore a pattern's meaning even though it would be advantageous to do so. On the other hand, when no semantic information was provided explicitly, the listeners may have simply constructed their own appropriate labels and frames for the patterns.

In conclusion, we have argued that many complex sound patterns have both syntactic and semantic structure which is determined by the sequence of source events which produce them. In interpreting such patterns, human listeners rely on their knowledge of these factors as well as on the perceptual information available in the sound itself. Most theorists agree that this occurs in the processing of linguistic information, and current research is underscoring the importance of syntactic and semantic factors in the perception of complex visual scenes (Biederman, 1980). Despite this, however, the role of these factors in the classification of nonlinguistic acoustic patterns has not been demonstrated previously. In the present study we have shown that these factors can play a significant role in even relatively simple classification tasks. Additional work is needed to elaborate their effects and to determine the influence of pattern structure on more traditional psychoacoustic measures such as the listener's ability to resolve individual pattern components (Watson & Kelly, in press).

References

Biederman, I. Human information processing of real-world scenes. Paper presented at the U.S. Army Research Institute Colloquium on Selected Topics in Behavioral Science Research, April, 1980.

Bregman, A.S. The formation of auditory streams. In J. Requin (Ed.), Attention and performance VII. Hillsdale, New Jersey: Erlbaum, 1978.

Cole, R.A., & Jakimik, J. Understanding speech: How words are heard. In G. Underwood (Ed.), Strategies of information processing. New York: Academic Press, 1978.

Cole, R.A., & Jakimik, J. A model of speech perception. In R.A. Cole (Ed.), Perception and production of fluent speech. Hillsdale, New Jersey: Erlbaum, 1980.

Marslen-Wilson, W.D. & Tyler, L.K. The temporal structure of spoken language understanding. Cognition, 1980, 8, 1-71.

Marslen-Wilson, W.D., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.

Minsky, M. A framework for representing knowledge. In P.H. Winston (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Reber, A.S. Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology, 1969, 81, 115-119.

Reber, A.S. Implicit learning of synthetic languages: The role of instructional set. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 88-94.

Reber, A.S., & Allen, R. Analogic and abstraction strategies in synthetic grammar learning: A functionalist interpretation. Cognition, 1978, 6, 189-221.

Reber, A.S., & Lewis, S. Implicit learning: An analysis of the form and structure of a body of tacit knowledge. Cognition, 1977, 5, 331-361.

Schank, R.C., & Abelson, R.P. Scripts, plans, goals and understanding. Hillsdale, New Jersey: Erlbaum, 1977.

Warren, R.M. Perceptual restoration of missing speech sounds. Science. 1970, 167, 392-393.

Watson, C.S., & Kelly, W.J. The role of stimulus uncertainty in

the discrimination of auditory patterns. In D.J. Getty & J.H. Howard, Jr. (Eds.), _Auditory and visual pattern recognition._ Hillsdale, N.J.: Erlbaum, in press.

455:J0:716:lab
80u455-234

OFFICE OF NAVAL RESEARCH

Code 455

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CDR Paul R. Chatelier
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Department of the Navy

Director
Engineering Psychology Programs
Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217 (5 cys)

Director
Undersea Technology
Code 220
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Electronics & Electromagnetics Technology
Code 250
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Naval Analysis Programs
Code 431
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Director
Physiology Program
Code 441
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Department of the Navy

Special Assistant for Marine Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Commanding Officer
ONR Eastern/Central Regional Office
ATTN: Dr. J. Lester
Building 114, Section D
666 Summer Street
Boston, MA 02210

Commanding Officer
ONR Branch Office
ATTN: Dr. C. Davis
536 South Clark Street
Chicago, IL 60605

Commanding Officer
ONR Western Regional Office
ATTN: Mr. R. Lawson
1030 East Green Street
Pasadena, CA 91106

Commanding Officer
ONR Western Regional Office
ATTN: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
American Embassy, Room A-407
APO San Francisco, CA 96503

Director
Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375 (6 cys)

Department of the Navy

Dr. Robert G. Smith
Office of the Chief of Naval Operations, OP987H
Personnel Logistics Plans
Washington, D.C. 20350

Dr. Jerry C. Lamb
Display Branch
Code TD112
Naval Underwater Systems Center
New London, CT 06320

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Human Factors Department
Code N215
Naval Training Equipment Center
Orlando, FL 32813

Dr. Alfred F. Smode
Training Analysis and Evaluation Group
Naval Training Equipment Center
Code N-OOT
Orlando, FL 32813

CDR R. Gibson
Bureau of Medicine & Surgery
Aerospace Psychology Branch
Code 513
Washington, D.C. 20372

CDR Robert Biersner
Naval Medical R&D Command
Code 44
Naval Medical Center
Bethesda, MD 20014

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD 20014

CDR Thomas Berghage
Naval Health Research Center
San Diego, CA 92152

Department of the Navy

Dr. George Moeller
Human Factors Engineering Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Dr. James McGrath, Code 311
Navy Personnel Research and Development Center
San Diego, CA 92152

Navy Personnel Research and Development Center
Management Support Department
Code 210
San Diego, CA 92152

CDR P. M. Curran
Code 604
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Dr. Lloyd Hitchcock
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Ronald A. Erickson
Human Factors Branch
Code 3194
Naval Weapons Center
China Lake, CA 93555

Human Factors Engineering Branch
Code 1226
Pacific Missile Test Center
Point Mugu, CA 93042

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Department of the Navy

Mr. H. Talkington
Ocean Engineering Department
Naval Ocean Systems Center
San Diego, CA 92152

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Mr. Warren Lewis
Human Engineering Branch
Code 8231
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Robert French
Naval Ocean Systems Center
San Diego, CA 92152

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Mr. Arnold Rubinstein
Naval Material Command
NAVMAT 08D22
Washington, D.C. 20360

Commander
Naval Air Systems Command
Human Factors Programs
NAVAIR 340F
Washington, D.C. 20361

Commander
Naval Air Systems Command
Crew Station Design,
NAVAIR 5313
Washington, D.C. 20361

Mr. Phillip Andrews
Naval Sea Systems Command
NAVSEA 0341
Washington, D.C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 4701
Washington, D.C. 20360

Department of the Navy

Human Factors Section
Systems Engineering Test
Directorate
U.S. Naval Air Test Center
Patuxent River, MD 20670

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

LCDR W. Moroney
Code 55MP
Naval Postgraduate School
Monterey, CA 93940

Mr. Merlin Malehorn
Office of the Chief of Naval
Operations (OP 115)
Washington, D.C. 20350

Department of the Army

Mr. J. Barber
HQS, Department of the Army
DAPE-MBR
Washington, D.C. 20310

Dr. Joseph Zeidner
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and
Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground, MD 21005

U.S. Army Aeromedical Research Lab
ATTN: CPT Gerald P. Krueger
Ft. Rucker, AL 36362

ARI Field Unit-USAREUR
ATTN: Library
C/O ODCSPER
HQ USAREUR & 7th Army
APO New York 09403

Department of the Air Force

U.S. Air Force Office of Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332

Dr. Donald A. Topmiller
Chief, Systems Engineering Branch
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Air University Library
Maxwell Air Force Base, AL 36112

Foreign Addressees

North East London Polytechnic
The Charles Myers Library
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

Professor Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich
Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Technology Establishment
Teddington, Middlesex TW11 OLN
ENGLAND

Director, Human Factors Wing
Defence & Civil Institute of Environmental Medicine
Post Office Box 2000
Downsview, Ontario M3M 3B9
CANADA

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF
ENGLAND

Other Government Agencies

Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 cys)

Dr. Craig Fields
Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Blvd
Arlington, VA 22209

Mr. M. Montemerlo
Human Factors & Simulation Division
NASA HQS
600 Independence Avenue
Washington, D.C. 20546

Dr. J. Miller
National Oceanic and Atmospheric Administration
11400 Rockville Pike
Rockville, MD 20852

Other Organizations

Dr. Robert R. Mackie
Human Factors Research, Inc.
5775 Dawson Avenue
Goleta, CA 93017

Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Dr. Robert G. Pachella
University of Michigan
Department of Psychology
Human Performance Center
330 Packard Road
Ann Arbor, MI 48104

Dr. Arthur I. Siegel
Applied Psychological Services, Inc.
404 East Lancaster Street
Wayne, PA 19087

Other Organizations

Dr. Gershon Weltman
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Williges
Human Factors Laboratory
Virginia Polytechnical Institute
and State University
130 Whittemore Hall
Blacksburg, VA 24061

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, NW.
Washington, D.C. 20036

Journal Supplement Abstract Service
American Psychological Association
1200 17th Street, N.W.
Washington, D.C. 20036 (3 cys)

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics
Company
St. Louis Division
Box 516
St. Louis, MO 63166

Dr. David J. Getty
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

DATE
FILMED

8